A Quasi-Copysafe Security of Documents on Normal Papersheets

Dr. Gyula Mezey

Associated Professor

Zrinyi Miklos National Defence University

(ZMNE VSZTK Vszt)

Address : 1581Budapest Pf.15 Hungary

Phone : (361)-3381971

E-mail : MEZEY@ZMNE.HU

Abstract

A combination of 2D barcode with digital signature and normal text with polygonal watermark is proposed. Against synchronisation attacks the watermark reference points are also included in the 2D barcode and secured by a digital signature, whilst the 2D barcode block(s) are embedded in the text.

Keywords : 2D barcodes, robust digital watermarks, digital signature, copy-management, access control

1. Security of documents

Security of documents is a general user requirement. During their lifecycle documents nowadays are born as electronic digital originals and printed only later. Copies of documents are transformations converting the same content either onto digital or analogue form. One factor of the security is the confidentiality of the document content. In this paper the confidentiality of a document content printed onto normal papersheet pages is in the focus. Let's assume that such a confidential system is going to be developed.

A basic system funtion : **Confidentiality**

Basic assumptions :

It is a general human habit, and so an implicit user requirement, that every important document is going to be printed.

Executives dislike watching a monitor screen, when reading over lengthy materials.

Our hypothesis is that this is thrue for confidential/secret materials as well, so it is a requirement to protect confidentiality/secrecy of printed documents by means of copy management and logical access control methods.

System functions :

Copy management

Logical access control

Techniques applicable :

Encryption systems

Steganography/digital watermarking

Access control systems

Access control : An access control system resist unauthorized access to the data

Encryption : Encryption resist unauthorized access to the content of a document.

Paper presented at the RTO SCI Symposium on "Strategies to Mitigate Obsolescence in Defense Systems Using Commercial Components", held in Budapest, Hungary, 23-25 October 2000, and published in RTO MP-072.

Steganography : Hiding/embedding the secret information beyond a cover text – that is steganograpy. But unlike encryption, steganograpy in itself does not resist access to the data and it is effective until the detecting of the hidden communication.

Problems : Used in point-to-point communication the information channel can be subject of various hostile attacks (jitter attack, etc.) aiming to fool the receiver/detector by either impairing or diminishing or removing the secret message.
After a successful hostile attack the hidden message cannot be recovered. As to paper media attacks these are usually detection-disabling or desynchronization geometric data manipulations.

2. Technical steganography approaches

Various techniques of technical steganography can be applied as well :
Even if the document content is encrypted, spread spectrum modulation, scattering make difficult to detect or jam transmission. Camouflage, special inks, materials, masking algorithms are widely in use as follows :

Blind colour approach

A blind colour is „invisible" for a copier/scanner/video/camera/etc. equipment.

Assumptions : An original document is printed by means of a colour printer, and
the really confidential/secret paragraphs, details, or data of the document are to be printed by a blind colour (red, orange, etc.), and
the rest of the confidential/secret document is to be printed by a non-blind colour.
Under the above assumptions the confidentiality of a photocopied/scanned/photod document – although copied – will be not corrupted indeed, because the confidential details are not on the copy.

Problems : Although a certain colour can be blind only for a subset of photocopier/scanner models, but not for all models.
Even with carefully selected photocopier models in office, original pages of confidential content can be fetched and photocopied/scanned outside by other models.
Video/photocamera, etc. can also smuggled in for copying the whole material.

Special copysafe papers are expensive and although having been copied on one model onto a copysafe special page the copy of the text is really unreadable, but having been copied on other colour copier models, the output can be a perfectly readable text.

Chemical reaction (heat or light effects) in the copier

Assumptions : The really confidential/secret paragraphs, details, or data of the document are to be printed by a special ink, and
in the printer there is no heat-effect when printing the original page, and
the background colour of the text is painted by a special ink (or the text on the page is printed by a special ink), wich will become of the same colour of the background, and
in the copier there is a heat-effect over a threshold temperature (or light effect), and
when making copies, only that copier is in use (heat-effect cannot be avoided.
Under the above assumptions the confidentiality of a photocopied/scanned/photod document – although copied – will be not corrupted indeed, because the confidential details are not on the copy.

Problems : There are too many underlying assumptions. In practice these prerequisits are difficult and in large organisations rather expensive to meet.

Conclusion : If we want to avoid risky as well as expensive proposals, we had better to find a commercial solution, perhaps a combination of normal paper and commercial ink and diverse printing equipment already existing in the environment. An integration of some carefully selected off-from the self commercial technologies will do a lot of good to the system.

3. Digital watermarking approaches

System subfunction : **Fingerprinting**

Copies of digital documents are indistinguishable from the original. Fingerprinting (hidden serial numbering) makes them distinguishable, which may become important in case of tracing for attackers. A hidden unique marking of each copy of a document makes distinguishable the original from the copies, and each copy from the other one. Hiding the secret information involves on the one hand hiding the location (or the reference) of the embedded information and on the other hand spreading the hidden information.

Digital watermarking techniques

On the expense of less embedded information into the cover text and using smarter methods, even if the communication has already been detected and the algorithmic principle of the embedding became public, a digital watermark can permanently reside in the host data, able to resist against hostile attacks, but unlike encryption, fragile watermarking in itself does not resist access to the data.

Robust digital watermarks are difficult to remove, impair from the cover text, but
fragile watermarks are relatively easy to remove, impair, destroy from the cover text.

Invisible watermarks

At least three invisible marking techniques are applicable for formatted black and white text printing:
By slightly changing

- interline spacing (Line-shift coding),
- intercharacter spacing (Word-shift coding),
- character font features (Character coding) a publisher can identify each document copy.

Because of underlying assumptions, line-shift decoding does not require the original unmarked copy. Although the marks robust enough to survive consecutive (ten generations) photocopying (see Ref. 1), theoretically removable, corruptable. The marks are fragile digital watermarks.
Problems : Any black and white marks in a formatted black and white textpage, layed out by any technique can always be removed by simply retyping, or by means of high resolution scanners and optical character recognition (OCR/ICR) and reprinting the text using a new character font and layout format.

Visible watermarks

In offices there has always been a requirement to register copies of documents. 1D barcode or serial numbers identified each copy (and pages of the respective copy) on the margin. By means of security printing scrambled barcode can be made in printing houses.

When watermarks generated, they should be innumerable (distinguishable from each other). Because of tracing back pirated copies, in fingerprinting applications it is important to identify the recipient of each individual distributed copy in the watermarks respectively. The identifier of the sender, the event (when and where) identifying attributes, and a document specific message digest are useful as well. In order to aviod taking a valid mark from one copy and pasting it onto another one, the sender signs the mark with a cryptographic key. Digital signature ensures, that the electronic document has not been altered.

In order to ensure security against hostile attacks (data manipulation), erasure of the watermark, or unauthorized access of the watermark content, cryptographic keys are very useful.
- Steganograpy in combination with a symmetric(secret) key – that is secret key watermarking.
- Steganograpy in combination with an asymmetric(public) key – that is public key watermarking.

Public key watermarking

A watermark is only robust as long as it is cannot be read by everyone. Public watermarks (where the key is public), are vulnerable to attacks unless each receiver uses a different key (but this is difficult in practice). Scrambled images can be descrambled by means of an optical grid or lens. A multilevel authentication system was proposed (see Ref . 2),where scrambled images can be verified by

variable (electooptical) filters, unique optical decoders. Another option is a special scrambling hardware in the camera.

Assumptions when copying a scrambled image, if the image is coloured (but not the text), and during printing a proper data resolution/frequency (1200 bpi) used, and a commercially available colour photocopier is applied, than the embedded mark cannot be copied.

Problems : But grayscale, or black and white scrambled images are copiable.

High resolution /non-commercial colour scanners can make perfect copies of the embedded images as well.

The reason why scrambled indicia is still used in offset or intaglio security printing (banknotes) and recently in personalisation of personal ID, is that the scrambling algorithm is secret (the key is in the hardware in the camera).

Secret key watermarking

Spread-spectrum radio communication is a symmetric key cryptosystem. The band spread is accomplished by a secret key(a signal, which is independent of the data) and a synchronized by the key reception at the receiver is used for despreading.

An example (see Ref. 3) to that, when parameters of a control signal (sequential impulse groups characteristics) are exploited to represent autonomous control information of remote control for vehicles or flying objects. That special data communication technique was proposed for transmission of some autonomous control attributes in a common one way channel parallel at the same time (Patent No.206418). The method in itself is independent of the field of application, circuit solution, or communication media. The communication channel can either be a cable, or a radiochannel (with different modulation options), either ultrasonic, or infrared, etc.

The modulating signal - an impulse group - can be seen in Fig. 1. One of the autonomous attributes is represented by the impulse-number in an impulse group (7). Another information is represented by the time interval between two starting pulses of two consecutive pulse groups (8). The third information is represented by the time interval rate of the existance (5) and nonexistance (6) of a certain pulse.

It is almost impossible to remove or replace a watermark, when that requires the secret key.

Problems : Attackers rather try to modify the watermark content. Or try to discredit the authority of the watermark by some ambiguity (inversion or interpretation) attack. By means of reverse engineering many watermarking schemes might be approximated.

Fake original documents, fake watermark data can be made. If different watermarks are embedded in the same host data it ought to be still possible to identify the first(authoritative or copyright) watermark.

Possible solutions : But also some methods are devised to construct noninvertible watermarks so that making them signal-dependent.

Information-losing marking schemes are also non-invertible, since inverses cannot be approximated closely enough.

A combination of watermarking and timestamping (provided by a trusted third party), or notarization, or a combination of watermarking, timestamping, cryptography, access control, 2D barcode.

4. Conditional logical access control

2D barcode approach

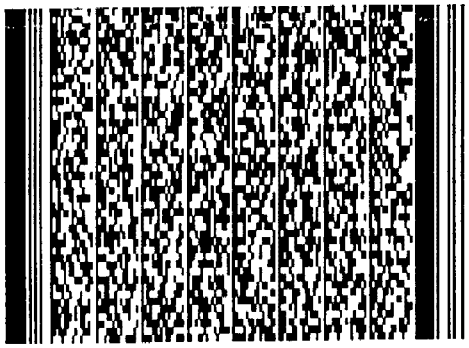
Assumptions : 2D barcode reader and printer,
and an organisationwide logical access control system are available, and
the really secret paragraphs, details, or data of the document are to be printed by 2D barcode (see Fig. 2) or glyph, and
the document qualification is : secret, than a public key encryption is also used, and
ciphering and deciphering software/hardware is also available, and
the rest of the confidential/secret document is to be printed as a normal text, than only the relatively short secret details (resolutions or important data in the protocol of an executives' meeting, etc.) ought to be watched by executives on a screen.

Under the above assumptions the secret content of a photocopied/scanned/photod document – although photocopied – will be not corrupted indeed, because though the confidential details are in the copy (and those remain copyable), but encryption (a digital signature or stamp) prevents the data content from unauthorized logical access. The latter has as good information protection performance as of a digital signature over a digital electronic document or file.

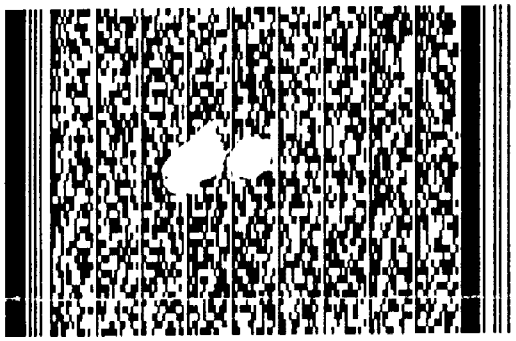
A combination of 2D barcode with digital signature (see Ref. 4) and normal text with polygonal watermark (see Ref. 5) is proposed. Against synchronisation attacks the watermark reference points are also included in the 2D barcode and secured by a digital signature, whilst the 2D barcode block(s) are embedded in the text.

REFERENCES

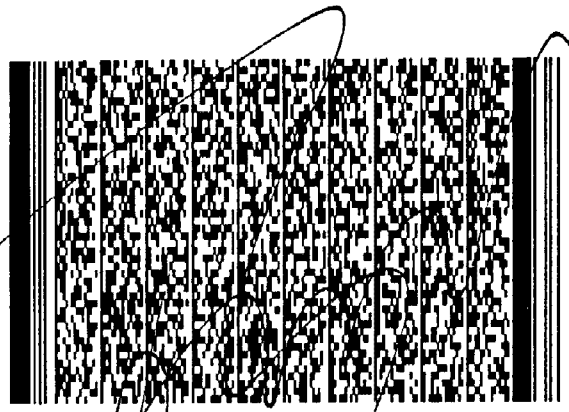
1. J.Brassil, S.Low, N.Maxemchuk, O'Gorman : Electronic Marking and Identification Techniques to Discourage Document Copying, IEEE J. Select. Areas Comm. , Vol. 13., pp.1495-1504. Oct.1995.
2. Mezey Gy.: A Multilevel Authentication System, In: P.Sugár, G.Chroust: CON'96 Future Information Technologies, Proc.11th Austrian-Hungarian Inf. Conf., Eger, Hungary 4-6. Nov.1996. pp.124-128.
3. -: A Special Data Communication Technique for Transmission of some Autonomous Control Attributes, In: Proc. Intl.Conf. on Advanced Military Technology in the XXI century – Concepts of the New Intelligence and Electronic Warfare Systems, Budapest, Hungary 16-17 May 2000. (ZMNE EK in press).
4. -: Digitális aláírás – papiron ? In: Proc.2st.Conf. Jogi Informatikai társaság (Law Inf.Soc.)/invited paper, in Hungarian/,Budapest, Hungary 18 March 2000
5. R. Obuchi et al. :Embedding Data in three-dimensional poligonal models, In : Proc. ACM Multimedia'97, Seattle, WA, Nov.1997



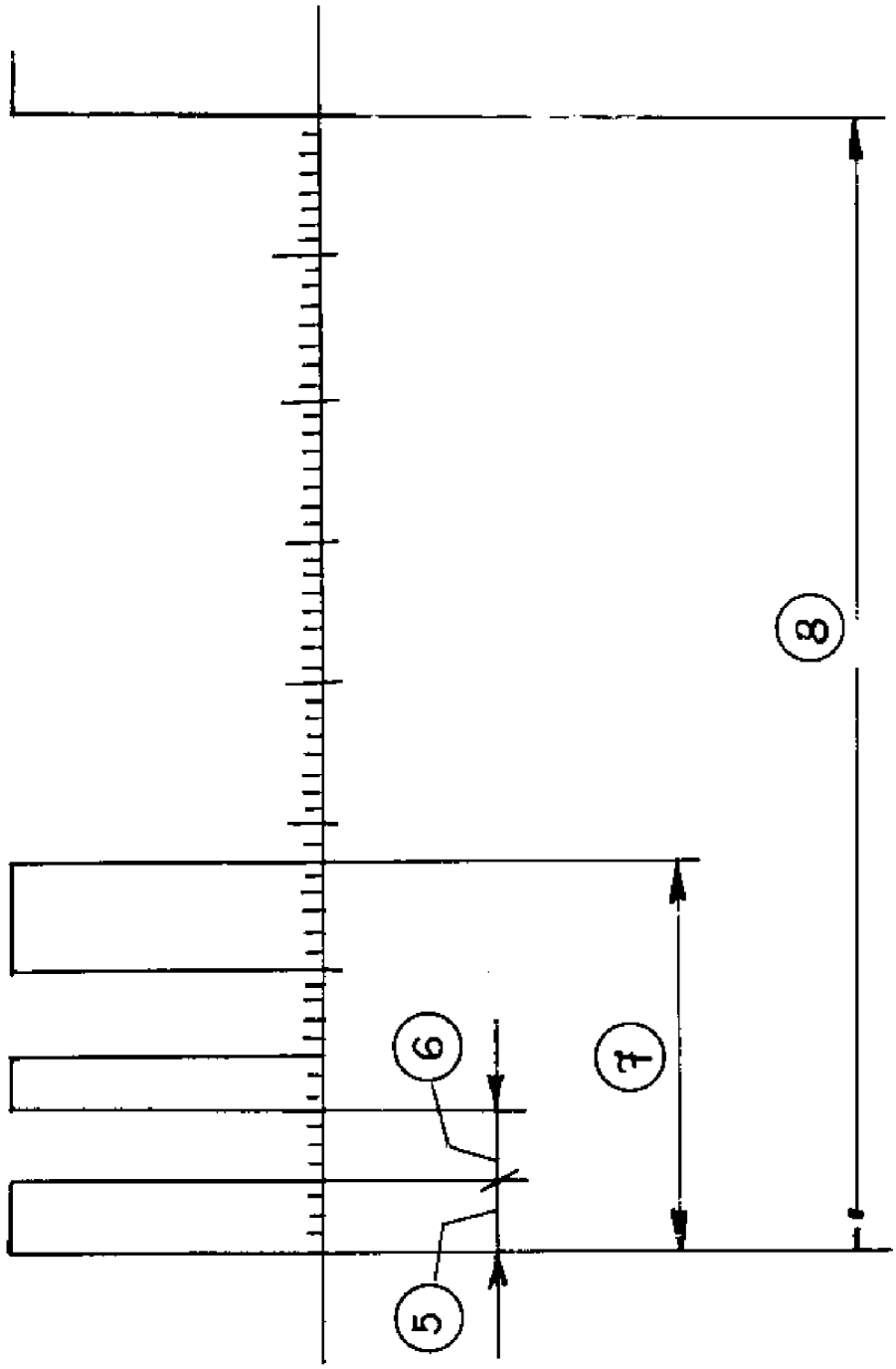
ECC: 10 %



ECC: 30 %



ECC: 50 %


8
7
6
5

This page has been deliberately left blank

Page intentionnellement blanche